

IN THE CLAIMS:

Please amend the claims to read as follows. This is a complete listing of all prior and pending claims and replaces any prior listing in this application.

1. (currently amended) **A data processing machine comprising a program storage device readable by said machine, said program storage device tangibly embodying a program of instructions executable by the machine to perform** a method for predicting the profitability of a commercial insurance policy, **said method** comprising:

~~gathering~~ **obtaining** policyholder data including premium and loss data ~~for storing in~~
from a database;

~~identifying external data sources~~ **obtaining external data** directed to at least one of business level data and household demographics data, the external data ~~sources~~ having a plurality of external variables to be used in predicting the profitability of the insurance policy;

associating the external variables with the policyholder data;

evaluating the associated external variables against the policyholder data to identify the individual external variables predictive of the insurance policy's profitability and creating a score based on an individually weighted multivariate statistical model based on said individual external predictive variables,

wherein said evaluating external variables includes evaluating the utility of creating new variables from the external variables and creating any appropriate new variables,

wherein said score is expressed as a sum of products, each of said products being a coefficient multiplied by a variable taken to a power,

and wherein said score is a function of at least all of the predictive external variables and any predictive new variables.

2. (cancelled)

3. (currently amended) The ~~method~~ **data processing machine** of claim 1 further comprising creating individual records in the database for each policyholder and populating each individual record with premium and loss data, business name, address and zip code for each policyholder and the associated external variables.

4. (currently amended) The ~~method~~ **data processing machine** of claim 1 further comprising associating at least one individual external variable with the individual records based on a unique data key associated with at least one external data source.

5. (currently amended) The ~~method~~ **data processing machine** of claim 1 further comprising normalizing the policyholder data in the database.

6. (currently amended) The ~~method~~ **data processing machine** of claim 5 wherein said normalizing further comprises premium manualization, loss trending and loss capping.

7. (original) The ~~method~~ **data processing machine** of claim 1 wherein the external data sources include external variables for at least one of geographic factors, business stability and weather patterns.

8. (currently amended) The ~~method~~ **data processing machine** of claim 1 wherein said evaluating the external variables further comprises examining the external variables for cross-correlation against one another in order to eliminate repetitive external variables.

9. (currently amended) The ~~method~~ **data processing machine** of claim 1 further comprising dividing the data in the database into a training data set for developing the statistical model, a testing data set for refining the statistical model and a validation data set for evaluating the predictiveness of the statistical model.

10. (currently amended) The ~~method~~ **data processing machine** of claim 1 wherein said identifying the external variables predictive of an insurance policy's profitability further includes normalizing the policyholder data, calculating for each policyholder the loss ratio based on the normalized policyholder data, defining a subgroup from the policyholder data, calculating a cumulative loss ratio for the subgroup and performing a statistical analysis to identify statistical relationships between individual external variables and the cumulative loss ratio for the subgroup.

11. (currently amended) The ~~method~~ **data processing machine** of claim 10 wherein the identified predictive external variables are examined for cross-correlations against one another.

12. (currently amended) The ~~method~~ **data processing machine** of claim 10 wherein the statistical model is created using multivariate methods to produce coefficients for each of the external predictive variables and the coefficients represent the contribution of the each of the external predictive variables to an overall score.

13. (currently amended) **A data processing machine comprising a program storage device readable by said machine, said program storage device tangibly embodying a program of instructions executable by the machine to perform** a method for creating a statistical model that generates a score representative of the profitability of an insurance policy for at least one of a new policyholder and an existing policyholder, **said method** comprising:

- gathering historical policyholder data, including loss and premium data;
- identifying external data sources having a plurality of external variables, each external variable having a value applying actuarial transformation to the policyholder data to generate working data;
- calculating a loss ratio for each policyholder in the database based on the working data;
- calculating a cumulative loss ratio for a defined group of policyholders in the database;
- performing a statistical analysis that investigates the relationship of each external variable and the cumulative loss ratio for the defined group to identify external variables that are predictive of the profitability of the insurance policy; and
- utilizing the predictive external variables identified in the previous step to develop an individually weighted multivariate statistical model that generates a score predictive of the profitability of the insurance policy,

wherein said performing a statistical analysis includes evaluating the utility of creating new variables from the external variables and creating any appropriate new variables,

wherein said score is expressed as a sum of products, each of said products being a coefficient multiplied by a variable taken to a power,

and wherein said score is a function of at least all of the predictive external variables and any predictive new variables.

14. (currently amended) The ~~method~~ **data processing machine** of claim 13 wherein the statistical model is used to score at least one of an existing policyholder and a new policyholder in order to determine the premium for a commercial insurance policy.

15. (currently amended) The ~~method~~ **data processing machine** of claim 13 further comprising manualizing the premium data, actuarially modifying long tail losses and capping large losses.

16. (currently amended) The ~~method~~ **data processing machine** of claim 13 further comprising binning together similar values of an external variable having multiple values.

17. (currently amended) The ~~method~~ **data processing machine** of claim 13 further comprising examining the external variables for cross-correlation against one another in order to eliminate repetitive external variables.

18. (currently amended) The ~~method~~ **data processing machine** of claim 13 further comprising dividing the data in the database into a training data set for developing the statistical model, a testing data set for refining the statistical model and a validation data set for evaluating the predictiveness of the statistical model.

19. (currently amended) The ~~method~~ **data processing machine** of claim 13 wherein the statistical model is created using multivariate methods to produce coefficients for each of the external predictive variables and wherein said coefficients represent the contribution of the each of the external predictive variables to the score.

20. (currently amended) A system for predicting the profitability of an insurance policy comprising:

a database for storing policyholder data including premium and loss data;

means for processing data from external data sources having a plurality of external variables to be used in predicting the profitability of the insurance policy;

means for evaluating the external variables against the policyholder data to identify the external variables predictive of an insurance policy's profitability; and

means for creating an individually weighted multivariate statistical model based on the external predictive variables that generates a score representative of the profitability of the insurance policy,

wherein said identifying external predictive variables includes evaluating the utility of creating new variables from the external variables and creating any appropriate new variables,

wherein said score is expressed as a sum of products, each of said products being a coefficient multiplied by a variable taken to a power,

and wherein said score is a function of at least all of the predictive external variables and any predictive new variables.

21. (previously presented) The system of claim 20 wherein the means for creating the individually weighted multivariate statistical model includes a software application for performing a multivariate statistical method on the external predictive variables to generate a coefficient for each external predictive variable, each coefficient representing the contribution of each external predictive variable to the score.

22. (original) The system of claim 21 wherein the multivariate method includes at least one of multiple regression and generalized linear modeling.

23. (currently amended) A system for creating an individually weighted multivariate statistical model that generates a score representative of the profitability of an insurance policy for at least one of a new policyholder and an existing policyholder, comprising:
a database of policyholder data
a means for accessing external data sources having multiple external variables;
predictive means for sorting through the external variables to identify individual external variables that are predictive of the profitability of the insurance policy; and
means for performing a statistical method on the external predictive variables to generate a coefficient for each external predictive variable, each coefficient representing the

contribution of each external predictive variable to the score generated by the statistical model,

wherein said predictive means includes means for evaluating the utility of creating new variables from the external variables and creating any appropriate new variables,

wherein said score is expressed as a sum of products, each of said products being a coefficient multiplied by a variable taken to a power,

and wherein said score is a function of at least all of the external predictive variables and any predictive new variables.

24. (previously presented) The system of claim 23 wherein the means for performing the statistical method comprises a software application that includes algorithms for performing at least one of multivariate statistical methods, clustering methods, decision tree techniques and neural network techniques.

25. (currently amended) **A data processing machine comprising a program storage device readable by said machine, said program storage device tangibly embodying a program of instructions executable by the machine to perform** a method of performing risk-based pricing of an insurance policy, comprising:
receiving a request for a price on an insurance policy; and
evaluating the risk associated with issuing the insurance policy based on a profitability score derived from an individually weighted multivariate statistical model generated with historical policyholder premium and loss data and external predictive variables identified from external

data sources, **said external data sources being** independent of internal policyholder data of an insurance company issuing the insurance policy,
wherein said identified external predictive variables include any predictive new variables created from external variables identified from said external data sources,
wherein said score is expressed as a sum of products, each of said products being a coefficient multiplied by a variable taken to a power,
and wherein said score is a function of at least all of the external predictive variables.

26. (currently amended) The ~~method~~ **data processing machine** of claim 25 wherein the external data sources include external variables for at least one of geographic factors, business stability and weather patterns.

27. (currently amended) The ~~method~~ **data processing machine** of claim 25 wherein the external data sources include at least one of business level data and household demographics data.

28. (currently amended) The ~~method~~ **data processing machine** of claim 25 further comprising examining the external predictive variables for cross-correlation against one another in order to eliminate repetitive external variables.

29. (currently amended) The ~~method~~ **data processing machine** of claim 25 wherein said identifying the external predictive variables further includes normalizing the policyholder data, calculating for each policyholder the loss ratio based on the normalized policyholder

data, defining a subgroup from the policyholder data, calculating a cumulative loss ratio for the subgroup and performing a statistical analysis to identify statistical relationships between individual external predictive variables and the cumulative loss ratio for the subgroup.

30. (currently amended) The ~~method~~ **data processing machine** of claim 28 wherein the identified external predictive variables are examined for cross-correlations against one another.

31. (currently amended) The ~~method~~ **data processing machine** of claim 25 wherein the statistical model is created using multivariate methods to produce coefficients for each of the external predictive variables and the coefficients represent the contribution of the each of the external predictive variables to an overall score.

32. (currently amended) The ~~method~~ **data processing machine** of claim 25 further comprising the step of dividing the policyholder data into a training data set for developing the statistical model, a testing data set for refining the statistical model and a validation data set for evaluating the predictiveness of the statistical model.

33. (currently amended) A **data processing** system for performing risk-based pricing of an insurance policy, comprising:

means for receiving a request for a price on an insurance policy; and

means for generating a profitability score derived from an individually weighted multivariate statistical model generated ~~with~~ **using** external variables identified from external data sources

independent of internal policy holder data of an insurance company issuing the insurance policy,

wherein said external variables include any new variables created from said external data sources,

wherein said score is expressed as a sum of products, each of said products being a coefficient multiplied by a variable taken to a power,

and wherein said score is a function of said external variables.

34. (original) The system of claim 33 wherein the external data sources include at least one of business level data and household demographics data.